Graph Cover Ensembles of Non-binary Protograph LDPC Codes

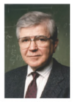
Dariush Divsalar   Lara Dolecek

JPL-Caltech   UCLA

ISIT 2012, MIT, Cambridge, MA

Motivation

- Gallager invented non-binary LDPC codes. Early work (MacKay and Davey, 1998) recognized the superiority of non-binary LDPC codes over binary LDPC codes.
- Majority of the subsequent results on graph-based non-binary codes were on assigning *non-zero* elements of Galois field on edges of final derived graph.


Related Work

- Non-binary LDPC codes
  - Lin, Abdel-Ghaffar, algebraic construction of non-binary LDPC codes .
  - Poulliat, Fossorier, Declercq 2008, Liva, Paolini, Scalise, Chiani, Costantini, Matuz 2011 and 2012.

Related Work

- Non-binary LDPC codes
  - Lin, Abdel-Ghaffar, algebraic construction of non-binary LDPC codes .
  - Poulliat, Fossorier, Declercq 2008, Liva, Paolini, Scalise, Chiani, Costantini, Matuz 2011 and 2012.
- Non-binary analysis
  - **Codewords**: Burshtein and Bennatan 2003, El-Khamy 2006, Kasai, Poulliat, Declercq 2008, Andriyanova 2009, Rosnes, Graell i Amat 2010, Savin, Declercq 2011, Divsalar and Dolecek 2011 and many others.
  - **Non-codewords (trapping sets and pseudocodewords)**: Kelley *et al.* 2006, Skachek and Flanagan 2008, Divsalar and Dolecek 2011.

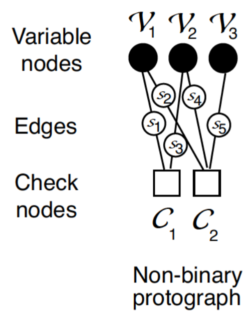
Related Work

- Non-binary LDPC codes
  - Lin, Abdel-Ghaffar, algebraic construction of non-binary LDPC codes .
  - Poulliat, Fossorier, Declercq 2008, Liva, Paolini, Scalise, Chiani, Costantini, Matuz 2011 and 2012.
- Non-binary analysis
  - **Codewords**: Burshtein and Bennatan 2003, El-Khamy 2006, Kasai, Poulliat, Declercq 2008, Andriyanova 2009, Rosnes, Graell i Amat 2010, Savin, Declercq 2011, Divsalar and Dolecek 2011 and many others.
  - **Non-codewords (trapping sets and pseudocodewords)**: Kelley *et al.* 2006, Skachek and Flanagan 2008, Divsalar and Dolecek 2011.

Our focus is on a new class of **structured non-binary** LDPC codes with graph cover construction.

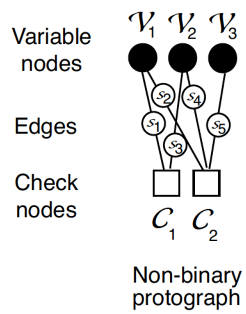
Graph Cover of Non-binary Protograph LDPC Codes

We introduce the *non-binary protograph* $G = (V, C, E, S)$ as a basic building block. It is a natural extension of binary protograph (Thorpe, 2003).



Graph Cover of Non-binary Protograph LDPC Codes

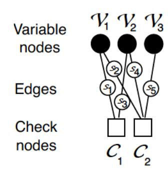
We introduce the *non-binary protograph* $G = (V, C, E, S)$ as a basic building block. It is a natural extension of binary protograph (Thorpe, 2003).



Edges are weighted by s_i's as non-zero elements of $GF(q)$.

Graph Cover of Non-binary Protograph LDPC Codes

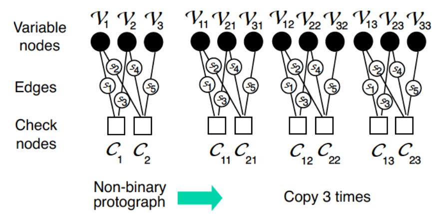
The protograph-based non-binary LDPC code is obtained by a copy-permute operation.



Non-binary protograph

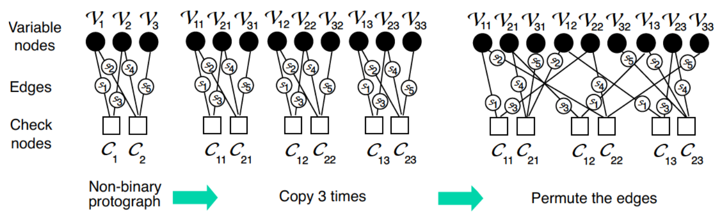
Graph Cover of Non-binary Protograph LDPC Codes

The protograph-based non-binary LDPC code is obtained by a copy-permute operation.



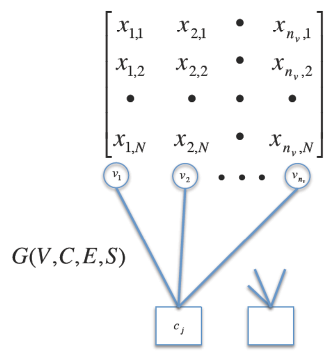
Graph Cover of Non-binary Protograph LDPC Codes

The protograph-based non-binary LDPC code is obtained by a copy-permute operation.

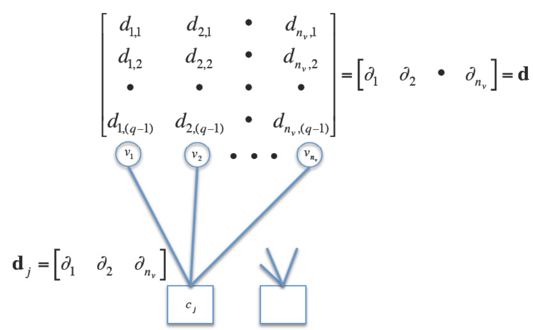


Codewords

Consider the *non-binary protograph* $G = (V, C, E, S)$ and degree-N graph cover of G as $G^{(N)}$. The codeword $\hat{\mathbf{x}}_N$ of $G^{(N)}$ can be represented as a matrix



Frequency weight matrix



The Hamming weight of a non-binary code is

$$d_H(\mathbf{d}(\hat{\mathbf{x}}_N)) = \sum_{i=1}^{n_v} \sum_{\ell=1}^{q-1} d_{i,\ell}. \quad (1)$$

Frequency weight enumerator of a check node

Notation

Frequency weight enumerator of a check node

Notation

- Let $\mathcal{C}_j$ be the single parity check code induced by check c_j of degree m_j.
- Let $K_j = q^{(m_j-1)}$ denote the number of codewords in $\mathcal{C}_j$.

Frequency weight enumerator of a check node

Notation

- Let $\mathcal{C}_j$ be the single parity check code induced by check c_j of degree m_j.
- Let $K_j = q^{(m_j-1)}$ denote the number of codewords in $\mathcal{C}_j$.
- Let $\mathbf{M}^{\mathcal{C}_j}$ be the $K_j \times m_j$ matrix with the codewords of $\mathcal{C}_j$ as its rows, and let $\mathbf{M}_b^{\mathcal{C}_j}$ be the $K_j \times m_j(q-1)$ *binary matrix* obtained as follows.

Frequency weight enumerator of a check node

Notation

- Let $\mathcal{C}_j$ be the single parity check code induced by check c_j of degree m_j.
- Let $K_j = q^{(m_j-1)}$ denote the number of codewords in $\mathcal{C}_j$.
- Let $\mathbf{M}^{\mathcal{C}_j}$ be the $K_j \times m_j$ matrix with the codewords of $\mathcal{C}_j$ as its rows, and let $\mathbf{M}_b^{\mathcal{C}_j}$ be the $K_j \times m_j(q-1)$ *binary matrix* obtained as follows.
- Consider a $1 \times m_j$ codeword $\mathbf{x} \in \mathcal{C}_j$. Let the mapping $\varphi(\mathbf{x})$ be defined as the symbol indicator, $\varphi(\mathbf{x}) =$ $[x_{1,1} \dots x_{1,(q-1)},\ x_{2,1} \dots x_{2,(q-1)}, \dots,\ x_{m_j,1} \dots x_{m_j,(q-1)}]$, where $x_{i\ell} = 1$, if the i-th component of $\mathbf{x}$ is equal to a non-binary symbol with index ℓ, otherwise $x_{i\ell} = 0$, for ℓ ranging over all $(q-1)$ non-zero symbols in $GF(q)$..

Frequency weight enumerator of a check node

It is convenient to view N replicas of a check node c_j of degree m_j with specified $\mathbf{s}_j$ as a $(m_j N, (m_j - 1)N)$ code $\mathcal{C}_j^N$ over $GF(q)$.

Theorem

The frequency weight matrix enumerator $A^{\mathcal{C}_j^N}(\mathbf{d}_j)$ of $\mathcal{C}_j^N$ is given by,

$$A^{\mathcal{C}_j^N}(\mathbf{d}_j) = \sum_{\{\mathbf{n}\}} C\left(N; n_1, n_2, \ldots, n_{K_j}\right), \tag{2}$$

where $C\left(N; n_1, n_2, \ldots, n_{K_j}\right)$ is the multinomial coefficient and $\{\mathbf{n}\}$ is the set of integer-vector solutions to $\mathbf{d}_j = \mathbf{n} \cdot \mathbf{M}_b^{\mathcal{C}_j}$, with $n_1, n_2, \ldots, n_{K_j} \geq 0$, and $\sum_{k=1}^{K_j} n_k = N$, and n_k the number of occurrences of the k^{th} codeword among these N copies of c_j.

Weight enumerator of a check node

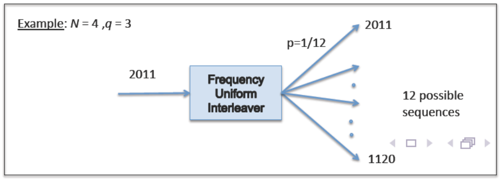
Proof (sketch)

- The weight-vector enumerator $A^{\mathcal{C}_j^N}(\mathbf{w})$ is derived as the coefficient of a multi-dimensional generating function of $\{A^{\mathcal{C}_j^N}(\mathbf{w})\}$.
- This generating function is expressed as the generating function of the code $\mathcal{C}_j$ (induced by the check node c_j and associated scale vector $\mathbf{s}_j$)), multiplied N times.
- An application of the multinomial theorem then provides the desired coefficients, which are then carefully collected to produce (2).

Frequency uniform interleaver (FUI)

Definition (Frequency uniform interleaver)

A length-N frequency uniform interleaver is a probabilistic device that maps each input of length N with entries as non-zero symbols of $GF(q)$ and with the frequency weight vector $[d_1, d_2, \cdots, d_{q-1}]$ into the $C(N; d_0, d_1, \ldots, d_{(q-1)})$ distinct output sequences of length N. Here $d_0 = N - \sum_{i>0} d_i$. These outputs have the same frequency weight vector as the input, and they are chosen equiprobably. ■



Weight enumerator of the GC-NBP ensemble

Let n_v be the number of variable nodes and n_c be the number of check nodes in G. Let t_i be the degree of variable node i.

Theorem

Let $A^{\mathcal{C}_j^N}(\mathbf{d}_j)$ be the frequency weight matrix enumerator of the code $\mathcal{C}_j^N$ induced by the N copies of the constraint node c_j with the associated scaling $\mathbf{s}_j$. Then, the frequency weight matrix enumerator of the GC-NBP ensemble is

$$A(\mathbf{d}) = \frac{\prod_{j=1}^{n_c} A^{\mathcal{C}_j^N}(\mathbf{d}_j)}{\prod_{i=1}^{n_v} C(N;\, d_{i,0},\, d_{i,1},\, \ldots,\, d_{i,(q-1)})^{t_i-1}}. \quad (3)$$

Weight enumerator of the GC-NBP ensemble

Proof.

Consider a concatenation of two codes, one induced by a variable node and another induced by a constraint node, connected via a frequency uniform interleaver. By collecting all the nodes, the result follows. □

- Finally, $\boxed{A_{d_H} = \sum_{\mathbf{d}} A(\mathbf{d}) \text{ where } \sum_{\mathbf{d}} d_{i,\ell} = d_H}$ is the average number of codewords of Hamming weight d_H in the GC-NBP ensemble.

Pairwise weight enumerators

- For NB protograph G create a modified NB protograph G^P.
- Each variable (constraint) node in G^P can be viewed as a duplicated version of a variable (constraint) node in the original protograph G.

Pairwise weight enumerators

- For NB protograph G create a modified NB protograph G^P.
- Each variable (constraint) node in G^P can be viewed as a duplicated version of a variable (constraint) node in the original protograph G.
- The local adjacency is preserved among the replicated nodes.

Pairwise weight enumerators

- For NB protograph G create a modified NB protograph G^P.
- Each variable (constraint) node in G^P can be viewed as a duplicated version of a variable (constraint) node in the original protograph G.
- The local adjacency is preserved among the replicated nodes.
- Each variable node in G^P takes a pair of elements of the field. This pair is interpreted as a new symbol (ℓ, ℓ') in the current set up..

Pairwise weight enumerators

- List all possible pairs of codewords of the original check node matrix $\mathbf{M}^{C_j}$ as the matrix $\mathbf{M}^{C_j}_{List}$
- The (i, k) component of the constraint matrix $\mathbf{M}^{C_j^P}$ of the new check node in G^P is the (i, k) and the $(i, k + m_j)$ component of $\mathbf{M}^{C_j}_{List}$ representing a new symbol (ℓ, ℓ') for $k = 1, \ldots, m_j$..

Pairwise weight enumerators

- List all possible pairs of codewords of the original check node matrix $\mathbf{M}^{C_j}$ as the matrix $\mathbf{M}_{List}^{C_j}$
- The (i,k) component of the constraint matrix $\mathbf{M}^{C_j^P}$ of the new check node in G^P is the (i,k) and the $(i,k+m_j)$ component of $\mathbf{M}_{List}^{C_j}$ representing a new symbol (ℓ,ℓ') for $k=1,\ldots,m_j$..
- The frequency weight vectors are defined as $\partial=[d_{0,0},\ldots,d_{(q-1),(q-1)}]^T$ where $d_{\ell,\ell'}$ counts the number of occurrences (frequency) of the symbol (ℓ,ℓ') within N symbols. The definition of the frequency weight matrix for protograph G^P is the same as previously discussed except that now the number of symbols is q^2 rather than q.

Pairwise weight enumerators

Theorem

The pairwise frequency weight matrix enumerator of the GC-NBP code averaged over the entire ensemble is

$$\tilde{A}(\mathbf{d}) = \frac{\prod_{j=1}^{n_c} A^{\mathcal{C}_{P,j}^N}(\mathbf{d}_j)}{\prod_{i=1}^{n_v} C(N;\ d_{i,0,0},\ d_{i,0,1},\ \ldots,\ d_{i,(q-1),(q-1)})^{t_i-1}},$$

where $A^{\mathcal{C}_{P,j}^N}(\mathbf{d}_j)$ is the frequency weight matrix enumerator of the code $\mathcal{C}_{P,j}^N$ induced by the N copies of the constraint node $c_{P,j}$ with the associated scaling $\mathbf{s}_j$. Here, the elements of $\mathbf{d}_j$ comprise a subset of the elements of $\mathbf{d} = [\partial_1, \partial_2, ..., \partial_{n_v}]$, and this subset (corresponds to neighbors of $c_{P,j}$) is obtained from the edge connections in the mother protograph G^P.

Pairwise weight enumerators

- For any two dimensional modulation with constellation points $a(\ell)$, we can define the pairwise Euclidean distance between two codewords as

$$d_E^2 = \sum_{i=1}^{n_v} \sum_{\ell=0}^{q-1} \sum_{\ell'=0}^{q-1} d_{i,\ell,\ell'} |a(\ell) - a(\ell')|^2. \qquad (4)$$

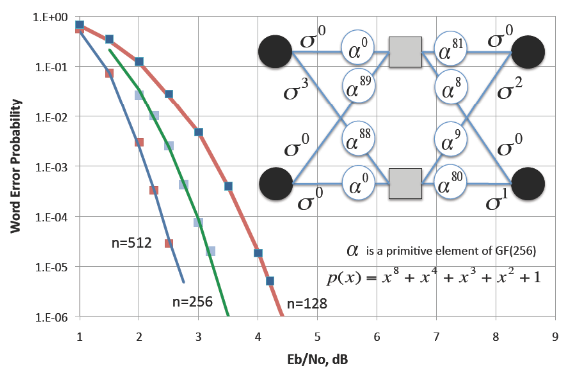
Pairwise weight enumerators

- For any two dimensional modulation with constellation points $a(\ell)$, we can define the pairwise Euclidean distance between two codewords as

$$d_E^2 = \sum_{i=1}^{n_v} \sum_{\ell=0}^{q-1} \sum_{\ell'=0}^{q-1} d_{i,\ell,\ell'} |a(\ell) - a(\ell')|^2. \qquad (4)$$

- Then $A_{d_E^2} = \sum_{\mathbf{d}} \tilde{A}(\mathbf{d})$ is the pairwise weight enumerator where the sum ranges over all frequency weight matrix $\mathbf{d}$ that produce channel dependent parameter d_E^2.

Example of GC-NBP codes



Conclusions and Future Work

Conclusions

- We introduced a new ensemble called *graph cover non-binary protograph LDPC codes*.
- Derivation of enumerators for codewords presented here. Results for trapping set, stopping set, pseudocodewords of GC-NBP codes and the asymptotic results are discussed in the paper.

On - going and future work

- Design and construction of practical GC-NBP and NB PB codes for short blocks [1].

[1] B. Y. Chang, D. Divsalar, and L. Dolecek, "Non-binary Protograph-Based LDPC Codes for Short Block-lengths," to appear in *IEEE ITW*, 2012.